DAIT Bioinformatics Summit 2009

Oct. 13/14, 2009
Rockville, Hilton
1750 Rockville Pike, Rockville, MD 20852

Registration: Bioinformatics Summit 2009 Registration Link

Background:

Integrating knowledge about disease phenotypes and finding preventive and treatment interventions present significant challenges for bioinformatics. During the 2006 Bioinformatics Summit meeting, we reviewed many programs funded by the National Institute of Allergy and Infectious Diseases (NIAID), Division of Allergy, Immunology and Transplantation (DAIT) and we identified several areas of potential collaborations, synergy and reusable resources within the different DAIT-funded bioinformatics and data management programs. In the 2007 Summit, we focused on the bioinformatics challenges of data reusability and secondary data analysis. In the 2008 Summit, we discussed the challenges of a systems approach for personalized medicine.

The main topics of the Summit 2009 are:

- Pros and cons of federated vs centralized approaches in managing translational research data
- Semantic web approaches for data integration
- Computational text mining and manual annotation in the development of biomedical knowledgebases
- Challenges of data visualization for a systems approach in biomedical research

Day 1 Oct. 13

12:00 pm Sign-in, Registration

1:00-1:15 **Opening Remarks**
Cheryl Kraft, Director, Office of Biomedical Informatics, DAIT

1:15- 2:15 **Keynote Presentation**
Measuring, Mining, Modeling, and Modulating: Translational Insights into the Acute Inflammatory Response
Yoram Vodovotz, Professor, Department of Surgery and Center for Inflammation and Regenerative Modeling, University of Pittsburgh

Session 1 Federated vs Centralized Approaches for Supporting Translational Research and Data Sharing

Topics:

- Evaluate the pros and cons of federated and centralized approaches
- Define the types of applications that are suitable for either a federated or centralized approach

Questions:

How does

- data diversity and complexity affect the decision to use either a federated or centralized approach?
- data volume affect the speed of data transfer, query and performance of an application in either approach?

What are the

- data security issues using federated or centralized approaches?
- technological complexities involved in performing queries in a federated model as opposed to a centralized mode?
- best criteria to evaluate the quality of data integration using either approach?

2:15- 2:45 **Architecture of an Adaptable Data Management System to Support Rapid Integration and Analysis of Scientific Data**
John Boyle, Director, Informatics at the Institute for Systems Biology

2: 45 – 3:00 **Break**

3:00-3:30 **Preventing the Incidentalome and Personalizing Medicine through Instrumentation of the Healthcare Enterprise**
Zak Kohane, Director, i2b2 National Center for Biomedical Computing

3:30- 4:00 **Data Unification for Integration: A Biomedical Data Processing and Integration Framework**
Xiaoming Wang, Computation Institute, University of Chicago and Argonne National laboratory

4:00 – 4:30 **Discussion**

Day 2 Oct. 14

Session 2 Challenges of Data Integration and Data Presentation for a Systems Approach in Translational Biomedical Research

Data integration and data visualization are two challenges of bioinformatics that go side-by-side. Each presentation contains both components with different emphasis.

Topics:

- Potential of semantic data integration with ontologies in data management, data visualization and data mining
- Potential and challenges of data capture, automated text mining and expert manual curation under a wiki framework to build publicly accessible knowledgebases.
- Creative query interfaces of biomedical repositories
- Creative data visualization applications to assist systems approaches in biomedical research
- Basic elements to build a user interface to meet the need of a systems approach in biomedical research

Questions:
What are the:

- Benefits and challenges of RDF based semantic data integration compared with a relational database approach with regard to ontology, development cycle, data storage, query capability and query performance?
- Challenges for community-based annotation of scientific knowledge using web 2.0 technology?
- Benefits and challenges of text mining tools that rely on automated natural language processing of clinical records and scientific literature?
- Key elements that ensure success in data integration for translational research?
- Challenges of developing data visualization tools to support translational research?
- Best methods for validating that these tools are integrated within normal workflow and provide the users the information they need?

7:30 am Sign-in, Registration

8:30 – 9:00 **Opportunities in Building Network Models of Biology and Applying Them to Human Disease**
Andrew Kasarskis, Sage Bionetworks and Fred Hutchinson Cancer Research Center

9:00 – 9:30 **Closing the Scientific Loop: Checking causality through Translational Knowledge Integration Using an Agent-based Modeling Format for Hypothesis Instantiation**
Gary An, Associate Professor of Surgery, Division of Trauma/Critical Care, Department of Surgery, Northwestern University

9: 30 – 10:00 **The Gene Wiki: Community Intelligence Applied to Gene Annotation**

Andrew I. Su, Head of the Computational Biology and Bioinformatics group, Genomics Institute of the Novartis Research Foundation

10:00 – 10:15 **Break**

10:15 - 10:45 **Information Visualization for Bioinformatics Discovery**
Ben Shneiderman, Professor, Department of Computer Science, University of Maryland

10: 45–11:15 **OntoMediator: A Query Integrator System Approach to Address Robust Data Integration at the Conceptual Level**
Luis N. Marenco, Associate Professor, Yale University Center for Medical Informatics

11:15 – 11:45 **Semantic Web Technology to Support Studying the Relation of HLA Structure Variation to Disease**
Alan Ruttenberg, Science Commons

11:45 – 12:15 **Discussion**

12:15 – 1:15 **Lunch on your own**

1:15 – 1:45 **The Creative Query Interfaces of Biomedical Repositories**,
Peter Cooper, NCBI

1:45-2:15 **Overcoming the Challenges to Discoverability in Healthcare and Life Science Research**
E. Sasha Paegle, Sr. Product Manager, Health Solutions Group, Microsoft

2:15 -2:45 **The Integration and Re-purposing of Clinical and Mechanistic Data from Multiple Clinical Trials: Creating a Reusable Resource for the Tolerance Community**
Dave Parrish, Executive Director, Informatics, Immune Tolerance Network

2:45-3:15 **Discussion**

3:15-3:30 **Closing Remarks**